

# Distance Functions for Reproducing Kernel Hilbert Spaces

N. Arcozzi, R. Rochberg, E. Sawyer, and B. D. Wick

**ABSTRACT.** Suppose  $H$  is a space of functions on  $X$ . If  $H$  is a Hilbert space with reproducing kernel then that structure of  $H$  can be used to build distance functions on  $X$ . We describe some of those and their interpretations and interrelations. We also present some computational properties and examples.

## 1. Introduction and Summary

If  $H$  is a Hilbert space with reproducing kernel then there is an associated set,  $X$ , and the elements of  $H$  are realized as functions on  $X$ . The space  $H$  can then be used to define distance functions on  $X$ . We will present several of these and discuss their interpretations, interrelations and properties. We find it particularly interesting that these ideas interface with so many other areas of mathematics.

Some of our computations and comments are new but many of the details presented here are known, although perhaps not as well known as they might be. One of our goals in this note is to bring these details together and place them in unified larger picture. The choices of specific topics however reflects the recent interests of the authors and some relevant topics get little or no mention.

The model cases for what we discuss are the hyperbolic and pseudohyperbolic distance functions on the unit disk  $\mathbb{D}$ . We recall that material in the next section. In the section after that we introduce definitions, notation, and some basic properties of Hilbert spaces with reproducing kernels. In Section 4 we introduce a function  $\delta$ , show that it is a distance function on  $X$ , and provide interpretations of it. In the section after that we introduce a pair of distance functions first considered in this context by Kobayashi and which, although not the same as  $\delta$ , are closely related. In Section 6 we discuss the relation between the distance functions that have been introduced and distances coming from having a Riemannian metric on  $X$ . The model case for this is the relation between three quantities on the disk, the

---

2010 *Mathematics Subject Classification.* 46E22.

*Key words and phrases.* Reproducing kernel, Hilbert space, metric, pseudohyperbolic.

The first author's work partially supported by the COFIN project Analisi Armonica, funded by the Italian Minister for Research.

The second author's work supported by the National Science Foundation under Grant No. 0700238.

The third author's work supported by the National Science and Engineering Council of Canada.

The fourth author's work supported by the National Science Foundation under Grant No. 1001098 and 0955432.

pseudohyperbolic distance, its infinitesimal version, the Poincare-Bergman metric tensor, and the associated geodesic distance, the hyperbolic metric.

Several of the themes here are common in recent literature on reproducing kernel Hilbert spaces. Some of the results here appear in the literature as results for Bergman spaces, but in hindsight they extend almost without change to larger classes of Hilbert space with reproducing kernel. Also, many results for the Hardy space suggest natural and productive questions for reproducing kernel Hilbert spaces with complete Nevanlinna Pick kernels. Those spaces have substantial additional structure and  $\delta$  then has additional interpretations and properties. That is discussed in Section 7.

Section 8 focuses on computations of how  $\delta$  changes when  $H$  is replaced by a subspace. Although that work is phrased in the language of distance functions it can be seen as yet another instance of using extremal functions for invariant or co-invariant subspaces as tools to study the subspaces.

Finally, we would like to emphasize that the material we present has not been studied much and most of the natural questions one can ask in this area are open.

## 2. Distances on the Unit Disk

Here we collect some background material; references are [G], [MPS], and [JP].

The pseudohyperbolic metric,  $\rho$ , is a metric on the unit disk,  $\mathbb{D}$ , defined by, for  $z, w \in \mathbb{D}$ ,

$$\rho(z, w) = \left| \frac{z - w}{1 - \bar{z}w} \right|.$$

Given any distance function  $\sigma$  we can define the length of a curve  $\gamma : [a, b] \rightarrow \mathbb{D}$  by

$$\ell_\sigma(\gamma) = \sup \left\{ \sum_{i=0}^{n-1} \sigma(\gamma(t_i), \gamma(t_{i+1})) : a = t_0 < t_1 < \dots < t_n = b \right\}.$$

Using this functional we can define a new distance,  $\sigma^*$ , by

$$\sigma^*(z, w) = \inf \{ \ell_\sigma(\gamma) : \gamma \text{ a curve joining } z \text{ to } w \}.$$

Automatically  $\sigma^* \geq \sigma$  and if equality holds  $\sigma$  is called an *inner distance*. More generally  $\sigma^*$  is referred to as the inner distance generated by  $\sigma$ .

The distance  $\rho$  is not an inner distance. The associated  $\rho^*$  is the hyperbolic distance,  $\beta$ , which is related to  $\rho$  by

$$\beta = \log \frac{1 + \rho}{1 - \rho}, \quad \rho = \frac{1}{2} \tanh \beta.$$

The hyperbolic distance can also be obtained by regarding the disk as a Riemannian manifold with length element

$$ds = \frac{2|dz|}{1 - |z|^2}$$

in which case  $\beta(z, w)$  is the length of the geodesic connecting  $z$  to  $w$ .

The Hardy space,  $H^2 = H^2(\mathbb{D})$ , is the Hilbert space of functions,  $f(z) = \sum_{n=0}^{\infty} a_n z^n$  which are holomorphic on the disk and for which  $\|f\|^2 = \sum |a_n|^2 < \infty$ . The inner product of  $f$  with  $g(z) = \sum b_n z^n$  is  $\langle f, g \rangle = \sum a_n \bar{b}_n$ . The Hardy space is a Hilbert space with reproducing kernel. That is, for each  $\zeta \in \mathbb{D}$  there is a kernel function  $k_\zeta \in H^2$  which reproduces the value of functions at  $\zeta$ ;  $\forall f \in H^2$ ,

$\langle f, k_\zeta \rangle = f(\zeta)$ . It is straightforward to see that there is at most one such function and that  $k_\zeta(z) = (1 - \bar{\zeta}z)^{-1}$  has the required property.

For the Hardy space now, and later for a general reproducing kernel Hilbert space, we are interested in the functional  $\delta(\cdot, \cdot)$ , defined for  $(z, w) \in \mathbb{D} \times \mathbb{D}$ , by

$$(2.1) \quad \delta(z, w) = \delta_{H^2}(z, w) = \sqrt{1 - \left| \left\langle \frac{k_z}{\|k_z\|}, \frac{k_w}{\|k_w\|} \right\rangle \right|^2}.$$

For the Hardy space this is evaluated as

$$(2.2) \quad \delta_{H^2}(z, w) = \sqrt{1 - \frac{(1 - |z|^2)(1 - |w|^2)}{|1 - \bar{z}w|^2}}.$$

This can be simplified using a wonderful identity. For  $z, w \in \mathbb{D}$

$$(2.3) \quad 1 - \frac{(1 - |z|^2)(1 - |w|^2)}{|1 - \bar{z}w|^2} = \left| \frac{z - w}{1 - \bar{z}w} \right|^2.$$

Hence  $\delta_{H^2}(z, w) = \rho(z, w)$ .

### 3. Reproducing Kernel Hilbert Spaces

By a reproducing kernel Hilbert space, RKHS, we mean a Hilbert space  $H$  of functions defined on a set  $X$  together with a function  $K(\cdot, \cdot)$  defined on  $X \times X$  with two properties; first,  $\forall x \in X$ ,  $k_x(\cdot) = K(\cdot, x) \in H$ , second  $\forall f \in H$   $\langle f, k_x \rangle = f(x)$ . The function  $k_x$  is called the reproducing kernel for the point  $x$ . We will use the following notation for unit vectors in the direction of the kernel functions. For  $x \in X$  we set

$$\hat{k}_x = \frac{k_x}{\|k_x\|}.$$

General background on such spaces can be found, for instance, in [AM]. Here we will just mention three families of examples and collect some standard facts.

#### 3.1. Examples.

##### The Dirichlet-Hardy-Bergman Family

For  $\alpha > 0$  let  $\mathcal{H}_\alpha$  be the RKHS of holomorphic functions on  $\mathbb{D}$  with the reproducing kernel

$$K_\alpha(w, z) = k_{\alpha, z}(w) = (1 - \bar{z}w)^{-\alpha}.$$

For  $\alpha = 0$  there is the limit version

$$K_0(w, z) = k_{0, z}(w) = \frac{1}{\bar{z}w} \log \frac{1}{1 - \bar{z}w}.$$

We have not included normalizing multiplicative constants as part of the kernels; we will be dealing only with expressions similar to (2.1) which are not affected by such constants. Also, we have not specified the norms for the spaces. In fact we will need to know how to take inner products with kernel functions but we will never need exact formulas for norms of general functions in the spaces. Hence we will give Hilbert space norms for the function spaces which are equivalent to the intrinsic RKHS norms.

First we consider the case  $\alpha > 1$ . These are generalized Bergman spaces;  $f \in \mathcal{H}_\alpha$  if and only if

$$\|f\|_{\mathcal{H}_\alpha}^2 \sim \int \int_{\mathbb{D}} |f(z)|^2 (1 - |z|^2)^{\alpha-2} dx dw < \infty.$$

The case  $\alpha = 2$  is the classical Bergman space.

If  $\alpha = 1$  we have the Hardy space described earlier. In that case the norm can be given using the radial boundary values  $f^*(e^{i\theta})$  by

$$\|f\|_{\mathcal{H}_0}^2 = \int_{\partial\mathbb{D}} |f^*(e^{i\theta})|^2 d\theta.$$

An equivalent norm for the Hardy space is

$$\|f\|_{\mathcal{H}_1}^2 \sim |f(0)|^2 + \int \int_{\mathbb{D}} |f'(z)|^2 (1 - |z|^2) dx dw.$$

The second description of the norm for the Hardy space is the one which generalizes to  $\alpha < 1$ . For  $0 \leq \alpha \leq 1$ ,  $f$  is in  $\mathcal{H}_\alpha$  exactly if

$$\|f\|_{\mathcal{H}_\alpha}^2 \sim |f(0)|^2 + \int \int_{\mathbb{D}} |f'(z)|^2 (1 - |z|^2)^\alpha dx dw < \infty.$$

The space  $\mathcal{H}_0$  is the Dirichlet space and the  $\mathcal{H}_\alpha$  for  $0 < \alpha < 1$  are called generalized Dirichlet spaces.

### The Fock-Segal-Bargmann Scale

For  $\beta > 0$  let  $F_\beta$  be the Hilbert space of holomorphic functions on  $\mathbb{C}$  for which

$$\|f\|_{F_\beta}^2 \sim \int \int_{\mathbb{C}} |f(z)|^2 e^{-\beta|z|^2} dx dy < \infty.$$

This is a RKHS and the kernel function is given by

$$K_\beta(z, w) = e^{\beta zw}.$$

**Remark 3.1.** There are other families of RKHS for which the kernel functions are powers of each other and still others where such relations hold asymptotically, see, for instance [E], [JPR].

### General Bergman Spaces

Suppose  $\Omega$  is a bounded domain in  $\mathbb{C}$  or, for that matter,  $\mathbb{C}^n$ . The Bergman space of  $\Omega$ ,  $B(\Omega)$ , is the space of all functions holomorphic on  $\Omega$  which are square integrable with respect to volume measure;  $f \in B(\Omega)$  exactly if

$$\|f\|_{B(\Omega)}^2 = \int \int_{\Omega} |f(z)|^2 dV(z) < \infty.$$

In this case it is easy to see that  $B(\Omega)$  is a Hilbert space and that evaluation at points of  $\Omega$  are continuous functionals and hence are given by inner products with some kernel functions. However, and this is one of the reasons for mentioning this example, it is generically not possible to write explicit formulas for the kernel functions.

**3.2. Multipliers.** Associated with a RKHS  $H$  is the space  $M(H)$  of multipliers of  $H$ , functions  $m$  defined on  $X$  with the property that multiplication by  $m$  is a bounded map of  $H$  into itself. For  $m \in M(H)$  we will denote the operator of multiplication by  $m$  by  $M_m$ . The *multiplier norm* of  $m$  is defined to be the operator norm of  $M_m$ .

For example, the multiplier algebra  $M(B(\Omega))$  consists of all bounded analytic functions on  $\Omega$ . The multiplier algebra of any of the spaces  $F_\beta$  consists of only the constant functions.

**3.3. Background Facts.** Suppose that  $H$  is a RKHS on  $X$  with kernel functions  $\{k_x\}_{x \in X}$  and multiplier algebra  $M(H)$ . The following are elementary Hilbert space results.

**Proposition 1.** Suppose  $f \in H$ ,  $\|f\| \leq 1$ ,  $x \in H$ . The maximum possible value of  $\operatorname{Re} f(z)$  (and hence also of  $|f(z)|$ ) is the value  $\|k_z\| = k_z(z)^{1/2}$  attained by the unique function  $f = \hat{k}_z$ .

**Proposition 2.** There is a unique  $F_{z,w} \in H$  with  $\|F_{z,w}\|_H \leq 1$  and  $F_{z,w}(z) = 0$  which maximizes  $\operatorname{Re} F_{z,w}(w)$ . It is given by

$$F_{z,w}(\cdot) = \frac{k_w(\cdot) - \|k_z\|^{-2} k_w(z) k_z(\cdot)}{\|k_w\| \sqrt{1 - |\langle \hat{k}_z, \hat{k}_w \rangle|^2}}.$$

and it has

$$F_{z,w}(w) = \|k_w\| \sqrt{1 - |\langle \hat{k}_z, \hat{k}_w \rangle|^2}.$$

**Proposition 3.** For  $m \in M(H)$ ,  $x \in X$  we have  $(M_m)^* k_x = \overline{m(x)} k_x$

**Proposition 4.** Suppose  $m \in M(H)$ ,  $x, y \in X$ . If  $\|M_m\|_{M(H)} \leq 1$  and  $m(x) = 0$  then

$$|m(y)| \leq \sqrt{1 - |\langle \hat{k}_x, \hat{k}_y \rangle|^2}.$$

#### 4. The Sine of the Angle

Suppose we have a RKHS  $H$  of functions on  $X$  and we want to introduce a metric on  $X$  that reflects the properties of functions in  $H$ . There are various ways to do this, for instance we could declare the distance between  $x, y \in X$  to be  $\|\hat{k}_x - \hat{k}_y\|$ . Here we focus on a different choice. Motivated by, among other things, the modulus of continuity estimates in Proposition 2 and Proposition 4 we define, if neither  $k_x$  nor  $k_y$  is the zero function,

$$(4.1) \quad \delta(x, y) = \delta_H(x, y) = \sqrt{1 - |\langle \hat{k}_x, \hat{k}_y \rangle|^2}.$$

We don't have a satisfactory definition of  $\delta(x, y)$  if  $k_x$  or  $k_y$  is the zero function. Either declaring these distances to be 1 or to be 0 would lead to awkwardness later. Instead we leave  $\delta$  undefined in such cases. However we will overlook that fact and, for instance write

$$\forall x, y \in X, \delta_{H_1}(x, y) = \delta_{H_2}(x, y)$$

to actually mean that the stated equality holds for all  $x, y$  for which both sides are defined.

One way to interpret  $\delta$  is to note that, by virtue of the propositions in the previous section,

$$\delta_H(x, y) = \frac{\sup \{|f(y)| : f \in H, \|f\| = 1, f(x) = 0\}}{\sup \{|f(y)| : f \in H, \|f\| = 1\}}.$$

Also,  $\delta(x, y)$  measures how close the unit vectors  $\hat{k}_x$  and  $\hat{k}_y$  are to being parallel. If  $\theta$  is the angle between the two then  $\delta(x, y) = \sqrt{1 - \cos^2 \theta} = |\sin \theta|$ .

In fact  $\delta$  is a pseudo-metric. It is clearly symmetric. It is positive semidefinite and will be positive definite if  $H$  separates points of  $X$ . (Although we will consider spaces which do not separate all pairs of points we will still refer to  $\delta$  as a metric.) The triangle inequality can be verified by a simple argument [AM, Pg. 128]. Instead we proceed to computations which develop further the idea that  $\delta$  measures the distance between points in the context of  $H$ . A corollary of the first of those computations is that  $\delta$  satisfies the triangle inequality.

For a linear operator  $L$  we denote the operator norm by  $\|L\|$  and the trace class norm by  $\|L\|_{\text{Trace}}$ . If  $L$  is a rank  $n$  self adjoint operator then it will have real eigenvalues  $\{\lambda_i\}_{i=1}^n$ . In that case we have

$$\|L\| = \sup \{|\lambda_i|\}, \quad \|L\|_{\text{Trace}} = \sum |\lambda_i|, \quad \text{Trace}(L) = \sum \lambda_i.$$

Also, recall that if  $L$  is acting on a finite dimensional space then  $\text{Trace}(L)$  equals the sum of the diagonal elements of any matrix which represents  $L$  with respect to an orthonormal basis.

**Proposition 5** (Coburn [CO2]). For  $x, y \in X$  let  $P_x$  and  $P_y$  be the self adjoint projections onto the span of  $k_x$  and  $k_y$  respectively. With this notation

$$\delta(x, y) = \|P_x - P_y\| = \frac{1}{2} \|P_x - P_y\|_{\text{Trace}}.$$

PROOF.  $P_x$  and  $P_y$  are rank one self adjoint projections and hence have trace one. Thus the difference,  $P_x - P_y$ , is a rank two self adjoint operator with trace zero and so it has two eigenvalues,  $\pm\lambda$  for some  $\lambda \geq 0$ . Thus  $\|P_x - P_y\| = \lambda$ ,  $\|P_x - P_y\|_{\text{Trace}} = 2\lambda$ . We will be finished if we show  $2\delta(x, y)^2 = 2\lambda^2$ . We compute

$$\begin{aligned} 2\lambda^2 &= \text{Trace}((P_x - P_y)^2) \\ &= \text{Trace}(P_x + P_y - P_x P_y - P_y P_x) \\ &= 2 - \text{Trace}(P_x P_y) - \text{Trace}(P_y P_x) \\ &= 2 - 2 \text{Trace}(P_x P_y). \end{aligned}$$

Going to last line we used the fact that  $\text{Trace}(AB) = \text{Trace}(BA)$  for any  $A, B$ . We now compute  $\text{Trace}(P_x P_y)$ . Let  $V$  be the span of  $k_x$  and  $k_y$ .  $P_x P_y$  maps  $V$  into itself and is identically zero on  $V^\perp$  hence we can evaluate the trace by regarding  $P_x P_y$  as an operator on  $V$ , picking an orthonormal basis for  $V$ , and summing the diagonal elements of the matrix representation of  $V$  with respect to that basis. We select the basis  $\hat{k}_y$  and  $j$  where  $j$  is any unit vector in  $V$  orthogonal to  $k_y$ . Noting that  $P_y \hat{k}_y = \hat{k}_y$  and  $P_y j = 0$  we compute

$$\begin{aligned} \text{Trace}(P_x P_y) &= \langle P_x P_y \hat{k}_y, \hat{k}_y \rangle + \langle P_x P_y j, j \rangle \\ &= \langle P_x \hat{k}_y, \hat{k}_y \rangle + 0 \\ &= \langle \langle \hat{k}_y, \hat{k}_x \rangle \hat{k}_x, \hat{k}_y \rangle \\ &= \left| \langle \hat{k}_y, \hat{k}_x \rangle \right|^2 \end{aligned}$$

which is what we needed.  $\square$

**Remark 4.1.** Because we actually found the eigenvalues of  $P_x$  and  $P_y$  we can also write  $\delta$  in terms of any of the Schatten  $p$ -norms,  $1 \leq p < \infty$ ;  $\delta(x, y) = 2^{-1/p} \|P_x - P_y\|_{S_p}$ .

A similar type of computation allows us to compute the operator norm of the commutator  $[P_a, P_b] = P_a P_b - P_b P_a$ . Informally, if  $a$  and  $b$  are very far apart,  $\delta(a, b) \sim 1$ , then each of the two products will be small and hence so will the commutator. If the points are very close,  $\delta(a, b) \sim 0$ , the individual products will be of moderate size and almost equal so their difference will be small.

**Proposition 6.**  $\|[P_a, P_b]\|^2 = \delta(a, b)^2 (1 - \delta(a, b)^2)$ .

PROOF. Note that  $[P_a, P_b]$  is a skew adjoint rank two operator of trace 0 and hence has eigenvalues  $\pm i\lambda$ , for some  $\lambda > 0$ . Hence  $\|[P_a, P_b]\| = \lambda$ . Also,  $[P_a, P_b]^* [P_a, P_b]$  is a positive rank two operator with eigenvalues  $\lambda^2, \lambda^2$  so its trace is  $2\lambda^2$ . We now compute

$$\begin{aligned} [P_a, P_b]^* [P_a, P_b] &= P_a P_b P_b P_a - P_a P_b P_a P_b - P_b P_a P_b P_a + P_b P_a P_a P_b \\ &= P_a P_b P_a - P_a P_b P_a P_b - P_b P_a P_b P_a + P_b P_a P_b \end{aligned}$$

Recalling that for any  $A, B$ ,  $\text{Trace}(AB) = \text{Trace}(BA)$  and also that the projections are idempotent we see that

$$\text{Trace}(P_a P_b P_a) = \text{Trace}(P_a P_b) = \text{Trace}(P_b P_a) = \text{Trace}(P_b P_a P_b).$$

The two middle quantities were computed in the previous proof

$$\text{Trace}(P_a P_b) = \text{Trace}(P_b P_a) = \left| \langle \hat{k}_a, \hat{k}_b \rangle \right|^2.$$

We also have

$$\text{Trace}(P_a P_b P_a P_b) = \text{Trace}(P_b P_a P_b P_a)$$

We compute the trace of the rank two operator  $P_b P_a P_b P_a$  by summing the diagonal entries of the matrix representation of the operator with respect to an orthonormal basis consisting of  $\hat{k}_a$  and  $j$ , a unit vector orthogonal to  $\hat{k}_a$ .

$$\begin{aligned} \text{Trace}(P_b P_a P_b P_a) &= \langle P_b P_a P_b P_a \hat{k}_a, \hat{k}_a \rangle + \langle P_b P_a P_b P_a j, j \rangle \\ (4.2) \quad &= \langle P_b P_a P_b \hat{k}_a, \hat{k}_a \rangle + 0 \end{aligned}$$

Next note that

$$\begin{aligned} P_b P_a P_b \hat{k}_a &= \langle \hat{k}_a, \hat{k}_b \rangle P_b P_a \hat{k}_b \\ &= \langle \hat{k}_a, \hat{k}_b \rangle \langle \hat{k}_b, \hat{k}_a \rangle P_b \hat{k}_a \\ &= \langle \hat{k}_a, \hat{k}_b \rangle \langle \hat{k}_b, \hat{k}_a \rangle \langle \hat{k}_a, \hat{k}_b \rangle \hat{k}_b \end{aligned}$$

and hence we can evaluate (4.2) and obtain  $|\langle \hat{k}_a, \hat{k}_b \rangle|^4$ . Thus

$$\begin{aligned} \| [P_a, P_b] \|^2 &= \frac{1}{2} \text{Trace} ([P_a, P_b]^* [P_a, P_b]) \\ &= \frac{1}{2} (2 \text{Trace} (P_a P_b) - 2 \text{Trace} (P_b P_a P_b P_a)) \\ &= |\langle \hat{k}_a, \hat{k}_b \rangle|^2 - |\langle \hat{k}_a, \hat{k}_b \rangle|^4 = \left(1 - |\langle \hat{k}_a, \hat{k}_b \rangle|^2\right) |\langle \hat{k}_a, \hat{k}_b \rangle|^2 \\ &= \delta(a, b)^2 (1 - \delta(a, b)^2). \end{aligned}$$

□

**Hankel forms:** The projection operators of the previous section can be written as  $P_a = \hat{k}_a \langle \hat{k}_a \rangle$ . In some contexts these are the primordial normalized Toeplitz operators. More precisely, on a Bergman space these are the Toeplitz operators with symbol given by the measure  $\|\hat{k}_a\|^{-2} \delta_a$  and general Toeplitz operators are obtained by integrating fields of these. There is also a map which takes functions on  $X$  to bilinear forms on  $H \times H$ . The basic building blocks for that construction are the norm one, rank one (small) Hankel forms given by  $L_a = \hat{k}_a \otimes \hat{k}_a$ , thus

$$L_a (f, g) = \langle f, \hat{k}_a \rangle \langle g, \hat{k}_a \rangle = f(a)g(a).$$

Limits of sums of these, or, equivalently, integrals of fields of these; are the Hankel forms on  $H$ ; for more on this see [JPR].

The norm of a bilinear form  $B$  on  $H \times H$  is

$$\|B\| = \sup \{ |B(f, g)| : f, g \in H, \|f\| = \|g\| = 1 \}.$$

Associated to a bounded  $B$  is a bounded conjugate linear map  $\beta$  of  $H$  to itself defined by  $\langle f, \beta g \rangle = B(f, g)$ . If we then define a conjugate linear  $\beta^*$  by  $\langle \beta^* f, g \rangle = \langle \beta g, f \rangle$  then

$$\langle \beta^* \beta f, f \rangle = \langle \beta f, \beta f \rangle \geq 0.$$

Thus  $\beta^* \beta$  is a positive linear operator. The form norm of  $B$  equals the operator norm of  $(\beta^* \beta)^{1/2}$  and we define the trace class norm of  $B$  to be the trace of the positive operator  $(\beta^* \beta)^{1/2}$ . With these definitions in hand we have a complete analog of Proposition 5.

**Proposition 7.** For  $x, y \in X$

$$\delta(x, y) = \|L_x - L_y\| = \frac{1}{2} \|L_x - L_y\|_{\text{Trace}}.$$

PROOF. Let  $\beta_x$  and  $\beta_y$  be the conjugate linear maps associated with  $L_x$  and  $L_y$ . One computes that  $\beta_x f = \beta_x^* f = \langle \hat{k}_a, f \rangle \hat{k}_a$  and similarly for  $\beta_y$ . Using this one then checks that for any  $x, y$ ;  $\beta_y^* \beta_x = P_y P_x$ . Thus the proof of Proposition 5 goes through. □

## 5. Formal Properties

We collect some observations on how the metric  $\delta$  interacts with some basic constructions on RKHS's.

**5.1. Direct Sums.** If  $H$  is a RKHS of functions on a set  $X$  and  $J$  is a RKHS on a disjoint set  $Y$  then we can define a RKHS  $(H, J)$  on  $X \cup Y$  to be the space of pairs  $(h, j)$  with  $h \in H, j \in J$  regarded as functions on  $X \cup Y$  via the prescription

$$(h, j)(z) = \begin{cases} h(z) & \text{if } z \in X \\ j(z) & \text{if } z \in Y. \end{cases}$$

One then easily checks that

$$\delta_{(H, J)}(z, z') = \begin{cases} \delta_H(z, z') & \text{if } z, z' \in X \\ \delta_J(z, z') & \text{if } z, z' \in Y \\ 1 & \text{otherwise} \end{cases}.$$

That computation is fundamentally tied to the fact that  $H$  and  $J$  are sets of functions on *different* spaces. If, however, all the spaces considered are functions on the *same* space then it is not clear what the general pattern is. That is, if  $H$  is a RKHS on  $X$  and if  $J, J'$  are two closed subspaces of  $H$  with, hoping to simplify the situation,  $J \perp J'$  then there seems to be no simple description of the relationship between  $\delta_H, \delta_J, \delta_{J'}$ , and  $\delta_{J \oplus J'}$ . In some of the examples in a later section we will compute these quantities with  $J' = J^\perp$  but no general statements are apparent.

**5.2. Rescaling.** Suppose  $H$  is a RKHS of functions on  $X$  and suppose that  $G(x)$  is a nonvanishing function on  $X$ ;  $G$  need not be in  $H$  and it need not be bounded. The associated rescaled space,  $GH$ , is the space of functions  $\{Gh : h \in H\}$  with the inner product

$$\langle Gf, Gg \rangle_{GH} = \langle f, g \rangle_H.$$

It is straightforward to check that  $GH$  is an RKHS and that its kernel function,  $K_{GH}$  is related to that of  $H$ ,  $K_H$  by

$$K_{GH}(x, y) = G(x)\overline{G(y)}K_H(x, y).$$

An immediate consequence of this is that  $\delta$  does not see the change;  $\delta_{GH} = \delta_H$ .

Elementary examples of rescaling show that certain types of information are not visible to  $\delta$ . Suppose we rescale a space  $H$  to the space  $cH$  for a number  $c$  (that is;  $|c|^2 \langle f, g \rangle_{cH} = \langle f, g \rangle_H$ ). The natural “identity” map from  $H$  to  $cH$  which takes the function  $f$  to the function  $f$  will be, depending on the size of  $c$ , a strict expansion of norms, an isometry, or a strict contraction. However it is not clear how one can recognize these distinctions by working with  $\delta_H$  and  $\delta_{cH}$ .

An awkward fact about rescaling is that sometimes it is present but not obviously so. Consider the following two pair of examples. First, let  $H$  be  $\mathcal{H}_1$ , the Hardy space of the disk. This can be realized as the closure of the polynomials with respect to the norm

$$\left\| \sum a_k z^k \right\|_{\mathcal{H}_1}^2 = \int_0^{2\pi} \left| \sum a_k e^{ik\theta} \right|^2 \frac{d\theta}{2\pi}.$$

For a weight, a smooth positive function  $w(\theta)$  defined on the circle, let  $\mathcal{H}_{1,w}$  be the weighted Hardy space; the space obtained by closing the polynomials using the norm

$$\left\| \sum a_k z^k \right\|_{\mathcal{H}_{1,w}}^2 = \int_0^{2\pi} \left| \sum a_k e^{ik\theta} \right|^2 w(\theta) \frac{d\theta}{2\pi}.$$

This is also a RKHS on the disk and in fact these two spaces are rescalings of each other. However to see that one needs to use a bit of function theory. The crucial fact is that one can write  $w(\theta)$  as

$$w(\theta) = |W(e^{i\theta})|^2$$

with  $W(z)$  and  $1/W(z)$  holomorphic in the disk and having continuous extensions to the closed disk. The functions  $W^{\pm 1}$  can then be used to construct the rescalings.

We now do a similar construction for the Bergman space. That space,  $\mathcal{H}_2$  in our earlier notation, is the space of holomorphic functions on the disk normed by

$$\|f\|_{\mathcal{H}_2}^2 = \int \int_{\mathbb{D}} |f(z)|^2 dx dy.$$

A weighted version of this space is given by replacing  $dx dy$  by  $w(z)dx dy$  for some smooth positive bounded  $w$ . To make the example computationally tractable we suppose  $w$  is radial;  $w(z) = v(|z|)$ . We define the weighted Bergman space  $\mathcal{H}_{2,w}$  by the norming function

$$\|f\|_{\mathcal{H}_{2,w}}^2 = \int \int_{\mathbb{D}} |f(z)|^2 w(z) dx dy.$$

The space  $\mathcal{H}_{2,w}$  is a RKHS on the disk and is an equivalent renorming of  $\mathcal{H}_2$  but is not related by rescaling. One way to see this is to note that, because the densities 1 and  $w(z)$  are both radial, in both cases the monomials are a complete orthogonal set. Thus, in both cases, the kernel function restricted to the diagonal is of the form

$$K(z, z) = \sum_0^{\infty} \frac{|z|^{2n}}{\|z^n\|^2} = a_0 + a_1 |z|^2 + \dots$$

Hence we can compute that for  $z$  near the origin

$$\delta(0, z) = \frac{\|1\|}{\|z\|} |z| (1 + O(|z|^2)).$$

If the spaces were rescalings of each other then the coefficients  $\|1\| / \|z\|$  would have to match, but this is not true independently of the choice of  $w$ .

**5.3. Products of Kernels.** In some cases the kernel function for a RKHS has a product structure. We begin by recalling two constructions that lead to that situation. Suppose that for  $i = 1, 2$ ;  $H_i$  is a RKHS on  $X_i$ . We can regard the Hilbert space tensor product  $H_1 \otimes H_2$  as a space of functions on the product  $X_1 \times X_2$  by identifying the elementary tensor  $h_1 \otimes h_2$  with the function on  $X_1 \times X_2$  whose value at  $(x_1, x_2)$  is  $h_1(x_1)h_2(x_2)$ . It is a standard fact that this identification gives  $H_1 \otimes H_2$  the structure of a RKHS on  $X_1 \times X_2$  and, denoting the three kernel functions by  $K_1$ ,  $K_2$ , and  $K_{1,2}$  we have

$$K_{1,2}((x_1, x_2), (x'_1, x'_2)) = K_1(x_1, x'_1) K_2(x_2, x'_2).$$

Now suppose further that  $X_1 = X_2$  and denote both by  $X$ . The mapping of  $x \in X$  to  $(x, x) \in X \times X$  lets us identify  $X$  with the diagonal  $D \subset X \times X$  and we will use this identification to describe a new RKHS,  $H_{12}$ , of functions on  $X$  (now identified with  $D$ ). The functions in  $H_{12}$  are exactly the functions obtained by restricting elements of  $H_1 \otimes H_2$  to  $D$ . The Hilbert space structure is given as follows. For

$f, g \in H_{12}$  let  $F$  and  $G$  be the unique functions in  $H_1 \otimes H_2$  which restrict to  $f$  and  $g$  and which have minimum norm subject to that condition. We then set

$$\langle f, g \rangle_{H_{12}} = \langle F, G \rangle_{H_1 \otimes H_2}.$$

To say the same thing in different words we map  $H_{12}$  into  $H_1 \otimes H_2$  by mapping each  $h$  to the unique element of  $H_1 \otimes H_2$  which restricts to  $h$  and which is orthogonal to all functions which vanish on  $D$ . The Hilbert space structure on  $H_{12}$  is defined by declaring that map to be a norm isometry. It is a classical fact about RKHSs that  $K_{12}$ , the kernel function for  $H_{12}$ , is given by

$$K_{12}(x, y) = K_1(x, y)K_2(x, y).$$

This leads to a relatively simple relation between the distance functions  $\delta_1$ ,  $\delta_2$  and  $\delta_{12}$  which we now compute. Pick  $x, y \in X$ . We have

$$\begin{aligned} 1 - \delta_{12}^2(x, y) &= \frac{|K_{12}(x, y)|^2}{K_{12}(x, x)K_{12}(y, y)} \\ &= \frac{|K_1(x, y)|^2}{K_1(x, x)K_1(y, y)} \frac{|K_2(x, y)|^2}{K_2(x, x)K_2(y, y)} \\ &= (1 - \delta_1^2(x, y))(1 - \delta_2^2(x, y)). \end{aligned}$$

Hence

$$\delta_{12} = \sqrt{\delta_1^2 + \delta_2^2 - \delta_1^2\delta_2^2}.$$

This implies the less precise, but more transparent, estimates

$$(5.1) \quad \max\{\delta_1, \delta_2\} \leq \delta_{12} \leq \delta_1 + \delta_2;$$

with equality only in degenerate cases. Similar results hold for  $\delta_{1,2}$ , the distance associated with  $H_1 \otimes H_2$ .

A particular case of the previous equation is  $H_1 = H_2$ . In that case

$$\delta_{12} = \sqrt{2\delta_1^2 - \delta_1^4} \geq \delta_1.$$

This monotonicity, which for instance relates the  $\delta$  associated with the Hardy space with that of the Bergman space, is a special case of the more general fact. If we have, for a set of positive  $\alpha$ , a family of spaces  $H_\alpha$  and if there is a fixed function  $K$  so that the kernel function  $K_\alpha$  for  $H_\alpha$  is  $K^\alpha$  then there is automatically a monotonicity for the distance functions; if  $\alpha' > \alpha$  then  $\delta_{\alpha'} \geq \delta_\alpha$ . This applies, for instance, to the families  $\{\mathcal{H}_\alpha\}$  and  $\{F_\beta\}$  introduced earlier. We also note in passing that in those two families of examples there is also a monotonicity of the spaces; if  $\alpha < \alpha'$  then there is a continuous, in fact a compact, inclusion of  $\mathcal{H}_\alpha$  into  $\mathcal{H}_{\alpha'}$ ; similarly for the  $F_\beta$ .

## 6. The Skwarczyński Metric

In [K] Kobayashi studies the differential geometry of bounded domains,  $\Omega$ , in  $\mathbb{C}^n$ . He begins with the observation that there was a natural map of  $\Omega$  into the projective space over the Bergman space of  $\Omega$ . He then notes that either of two naturally occurring metrics on that projective space could then be pulled back to  $\Omega$  where they would be useful tools. However looking back at his paper there was no particular use made of the Bergman space beyond the fact that it was a RKHS. We will now describe his constructions and see that they give expressions which are not the same as  $\delta$  but are closely related.

Suppose  $H$  is a RKHS of functions on  $X$ . Canonically associated with any point  $x \in X$  is a one dimensional subspace of  $H$ , the span of the kernel function  $k_x$ , or, what is the same thing, the orthogonal complement of the space of functions in  $H$  which vanish at  $x$ . The projective space over  $H$ ,  $P(H)$ , is the space of one dimensional subspaces of  $H$ . Hence for each  $x \in X$  we can use the span of  $k_x$ ,  $[k_x]$  to associate to  $x$  a point  $p_x = [k_x] \in P(H)$ . To understand the geometry of this mapping we break it into two steps. First, we associate to each  $x \in X$  the set of vectors in the unit sphere,  $S(H)$  that are in the span of  $k_x$ ; all these vectors of the form  $e^{i\theta} \hat{k}_x$  for real  $\theta$ . The next step is to collapse this circle sitting in the unit spheres,  $\{e^{i\theta} \hat{k}_x : \theta \in \mathbb{R}\}$ , to the single point  $p_x = [k_x]$ . In fact every point  $p \in P(H)$  is associated in this way to a circle  $C(p) \subset S(H)$  and distinct points correspond to disjoint circles. We now use the fact the distance function of  $H$  makes  $S(H)$  a metric space and use that metric to put the quotient metric on  $P(H)$ . That is, define a metric  $\hat{\delta}$  on  $P(H)$ , sometimes called the Cayley metric, by:

$$\hat{\delta}(p, q) = \inf \{\|r - s\| : r \in C(p), s \in C(q)\}$$

On the subset  $\{p_x : x \in X\}$  we have explicit descriptions of the circles  $C(p_x)$  and we compute

$$\begin{aligned} \hat{\delta}(p_x, p_y) &= \inf \left\{ \|e^{i\theta} \hat{k}_x - e^{i\eta} \hat{k}_y\| : \theta, \eta \in \mathbb{R} \right\} \\ &= \inf \sqrt{2 - 2 \operatorname{Re} e^{i(\theta-\eta)} \langle \hat{k}_x, \hat{k}_y \rangle} \\ &= \sqrt{2} \sqrt{1 - |\langle \hat{k}_x, \hat{k}_y \rangle|}. \end{aligned}$$

We now pull this metric back to  $X$  and, with slight abuse of notation, continue to call it  $\hat{\delta}$ ;

$$(6.1) \quad \hat{\delta}(x, y) = \sqrt{2} \sqrt{1 - |\langle \hat{k}_x, \hat{k}_y \rangle|}$$

Thus the map of  $X$  into  $P(H)$  which sends  $x$  to  $p_x$  is a isometry  $X$  with the metric  $\hat{\delta}$  into  $P(H)$  with its metric as a quotient of  $S(H)$ .

This metric was studied further by Skwarczyński in [S] and then in collaboration with Mazur and Pflug in [MPS]. In [JP] it is referred to as the Skwarczyński metric.

The second metric Kobayashi introduces in this context is again obtained by putting a natural metric on  $P(H)$  and then, again, pulling it back to  $X$ . The Fubini-Study metric is a natural Kahler metric in finite dimensional projective space and Kobayashi extends that definition to the generally infinite dimensional  $P(H)$ . We denote by  $\check{\delta}$  the metric obtained by restricting the Fubini-Study metric to the image of  $X$  in  $P(H)$  and then pulling the metric back to  $X$ .

In the small these three metrics are almost the same. If one is small so are the others and, in fact, setting  $\delta(x, y) = \delta_1$ ,  $\hat{\delta}(x, y) = \delta_2$ ,  $\check{\delta}(x, y) = \delta_3$  we have that

$$(6.2) \quad \max_{i=1,2,3} \{\delta_i\} = O(\varepsilon) \implies \max_{i,j=1,2,3} \{|\delta_i - \delta_j|\} = O(\varepsilon^3).$$

The comparison between  $\delta_1$  and  $\delta_2$  follows from (4.1), (6.1) and Taylor's theorem. The comparison of  $\delta_2$  and  $\delta_3$ , which is not difficult, is given in [K, pg. 282].

Comparing (4.1) and (6.1) also allows one to see that

$$\lim_{\delta(x,y) \rightarrow 1} \frac{\delta(x,y)}{\hat{\delta}(x,y)/\sqrt{2}} = 1$$

**Remark 6.1.** In mathematical physics, for certain choices of  $H$ , the map from  $X$  into the projective space is related to coherent state quantization. In that context some of the quantities we have been working with, or will be below, are given physical names/interpretations. For instance  $|\langle \hat{k}_x, \hat{k}_y \rangle|^2 = 1 - \delta(x,y)^2$  is the probability density for transition from the quantum state  $[k_x]$  to the state  $[k_y]$ . See, for instance, [O1], [O2], [AE], and [PV].

## 7. Differential Geometric Metrics

In this section we describe the relationship between the distance functions we introduced  $\delta$ , the associated length functions and inner metrics, and a Riemannian metric built using the kernel functions. Throughout this section we suppose that  $H$  is a RKHS of holomorphic functions on a domain  $X$  in  $\mathbb{C}$ . We further suppose that  $H$  is nondegenerate in the sense that  $\forall x, y \in X, \exists h, k \in H$  with  $h(x) \neq 0, k(x) \neq k(x)$ . These restrictions are much more than is needed for most of what follows but some restrictions are necessary. For instance the results in the next subsection require that the kernel function  $K(x, y)$  be sufficiently smooth so that one can apply to it the second order Taylor's theorem; some of the results in the third subsection are specific to one complex variable.

**7.1. Results of Mazur, Pflug, and Skwarczyński.** In an earlier section we described how, for distance functions on the unit disk, one could pass from a distance to the associated inner distance. That discussion was not specific to the disk and we now apply those constructions to distances defined on  $X$ . That is, given a distance function  $\mathcal{D}$  we define the length of a curve  $\gamma$  by  $\ell_{\mathcal{D}}(\gamma) = \sup \{ \sum \mathcal{D}(\gamma(t_i)\gamma(t_{i+1})) \}$  and the inner distance induced by  $\mathcal{D}$  is given by  $\mathcal{D}^*(x, y) = \inf \{ \ell_{\mathcal{D}}(\gamma) : \gamma \text{ is a curve from } x \text{ to } y \}$ . Clearly  $\mathcal{D}^* \geq \mathcal{D}$  and if the two functions are equal we say  $\mathcal{D}$  is an inner distance. For example Euclidean distance on the plane is an inner distance; the pseudohyperbolic distance in the disk is not an inner distance, its induced inner distance is the hyperbolic distance.

The reproducing kernel for the Bergman space of the disk,  $\mathcal{H}_2 = B(\mathbb{D})$ , the *Bergman kernel* is  $K(x, y) = (1 - xy)^{-2}$ . Using it we can construct a Riemannian metric on the disk through

$$\begin{aligned} ds^2 &= \frac{\partial^2}{\partial z \partial \bar{z}} \log k_z(z) |dz|^2 = \frac{1}{4} \Delta \log k_z(z) |dz|^2 \\ &= \frac{2}{(1 - |z|^2)^2} |dz|^2 \end{aligned}$$

which is a constant multiple of the density we saw earlier when introducing the hyperbolic metric on the disk.

More generally, if  $X$  is a bounded domain in  $\mathbb{C}$  and  $H$  is the Bergman space of  $X$ ,  $H = B(X)$ , and  $\{k_z\}$  are the reproducing kernels then the formula  $ds^2 = \partial \bar{\partial} \log k_z(z) |dz|^2$  defines a Riemannian metric on  $X$ , the so called Bergman metric. There is also an extension of this construction to domains in  $\mathbb{C}^n$ .

In [MPS] Mazur, Pflug, and Skwarczyński prove three theorems. Suppose that  $X$  is a bounded domain in  $\mathbb{C}$  (they actually work with  $\mathbb{C}^n$ ). Let  $H$  be the Bergman space of  $X$ . For a curve  $\gamma$  in  $X$  let  $\ell_B(\gamma)$  be its length measured using the Bergman metric. For  $x, y \in X$  denote the Bergman distance between them by  $\delta_B(x, y) = \inf \{\ell_B(\gamma) : \gamma \text{ a smooth curve from } x \text{ to } y\}$ .

**Proposition 8** ([K], [MPS]).

(1) For any smooth curve  $\gamma$

$$\ell_\delta(\gamma) = \ell_{\hat{\delta}}(\gamma) = \ell_{\check{\delta}}(\gamma) = \frac{1}{2}\ell_B(\gamma).$$

(2)

$$\delta^* = \hat{\delta}^* = \check{\delta}^* = \check{\delta} = \frac{1}{2}\delta_B.$$

(3)

$$\forall x, y \in X, x \neq y, \hat{\delta}(x, y) < \hat{\delta}^*(x, y).$$

In particular (up to a constant factor) the Bergman distance is the inner distance generated by our metrics and the metric  $\hat{\delta}$  is never an inner metric.

The results for  $\hat{\delta}$  are proved in [MPS] (with an unfortunate typo in the statement of Theorem 2 there). We noted that locally the three distances in (6.2) agree to third order. Hence the three metrics generate the same length function and same inner distance. Because of this the results for  $\delta$  and  $\check{\delta}$  follow from the ones for  $\hat{\delta}$ . The equality  $\check{\delta}^* = \check{\delta}$ , i.e. the statement that  $\check{\delta}$  is an inner distance, follows from the discussion in [K].

The discussion in [K] [MPS] is given for Bergman spaces,  $B(\Omega)$ . However the results hold in more generality. Given  $X$  and  $H$  of the type we are considering, with reproducing kernels  $\{k_z\}$ , there is a standard associated Riemannian metric given by

$$(7.1) \quad ds_H^2 = \left( \frac{\partial^2}{\partial z \partial \bar{z}} \log k_z(z) \right) |dz|^2.$$

If we define functions  $k_z^{(1)}$  in  $H$  by requiring that for all  $f \in H$   $\langle f, k_z^{(1)} \rangle = f'(z)$  then one can compute that

$$ds_H^2 = \frac{\|k_z^{(1)}\|^2 \|k_z\|^2 - \left| \langle k_z^{(1)}, k_z \rangle \right|^2}{\|k_z\|^4} |dz|^2.$$

One can then define the Bergman style length of a curve  $\gamma$ ,  $\ell_{BS}(\gamma)$ , to be the length of  $\gamma$  measured using  $ds_H$  and can set

$$\delta_{BS}(x, y) = \inf \{\ell_{BS}(\gamma) : \gamma \text{ a smooth curve connecting } x \text{ to } y\}.$$

We have defined  $\delta = \delta_H$  for such an  $H$ . We define  $\hat{\delta} = \hat{\delta}_H$  using (6.1). We define  $\check{\delta} = \check{\delta}_H$  by following Kobayashi's prescription. We have a map of  $X$  into the  $P(H)$  which sends  $x$  to  $p_x = [k_x]$ . We use that map to we pull back the Fubini-Study metric on  $P(H)$  back to  $X$  and call the resulting metric  $\check{\delta}$ .

**Proposition 9.**

(1) For any smooth curve  $\gamma$

$$\ell_\delta(\gamma) = \ell_{\hat{\delta}}(\gamma) = \ell_{\check{\delta}}(\gamma) = \frac{1}{2}\ell_{BS}(\gamma).$$

(2)

$$\delta^* = \hat{\delta}^* = \check{\delta}^* = \check{\delta} = \frac{1}{2}\delta_{BS}.$$

(3)

$$\forall x, y \in X, x \neq y, \hat{\delta}(x, y) < \hat{\delta}^*(x, y).$$

(4)

$$\forall x, y \in X, x \neq y, \delta(x, y) < \delta^*(x, y).$$

PROOF. The proof in [MPS] of versions of the first two statements are based on the second order Taylor approximations to the kernel functions; hence those proofs apply here as does the discussion in [K] which shows that  $\check{\delta}$  is an inner metric. The third statement follows from the proof in [MPS] together with the fact that for any  $a, b, c \in X$  we have the strict inequality  $\hat{\delta}(a, c) < \hat{\delta}(a, b) + \hat{\delta}(b, c)$ . The proof of that in [MPS] does not use the fact that  $H$  is a Bergman space, rather it uses the fact that  $\hat{\delta}$  was obtained by pulling the Cayley metric back from projective space, which also holds in our context. The fourth statement also follows from the proof in [MPS] if we can establish the fact that for any  $a, b, c \in X$  we have the strict inequality  $\delta(a, c) < \delta(a, b) + \delta(b, c)$ . We will obtain that from Proposition 5. We need to rule out the possibility that

$$(7.2) \quad \|P_a - P_c\| = \|P_a - P_b\| + \|P_b - P_c\|.$$

The operator  $P_a - P_c$  is a rank two self adjoint operator. Hence it has a unit eigenvector,  $v$ , with

$$\|(P_a - P_c)(v)\| = \|P_a - P_c\|.$$

For both of the two previous equalities to hold we must also have

$$\|(P_a - P_b)(v)\| = \|P_a - P_b\|, \quad \|(P_b - P_c)(v)\| = \|P_b - P_c\|$$

and hence  $v$  must also be an eigenvector of  $P_a - P_b$  and also of  $P_b - P_c$ .

However in our analysis in Proposition 5 we saw that an eigenvector for an operator  $P_x - P_y$  must be in  $\bigvee \{k_x, k_y\}$ , the span of  $k_x$  and  $k_y$ . Thus

$$v \in \bigvee \{k_a, k_c\} \cap \bigvee \{k_a, k_b\} \cap \bigvee \{k_b, k_c\} = \{0\},$$

a contradiction.  $\square$

**7.2. The Berezin Transform and Lipschitz Estimates.** Suppose  $A$  is a bounded linear map of  $H$  to itself. The Berezin transform of  $A$  is the scalar function defined on  $X$  by the formula

$$\hat{A}(x) = \langle A\hat{k}_x, \hat{k}_x \rangle.$$

For example, if  $P_a$  is the orthogonal projection onto the span of  $k_a$  then

$$\widehat{P}_a(x) = 1 - \delta^2(a, x).$$

Also, recalling Proposition 3, we have the following. Suppose  $m, n \in M(H)$  and that  $M$  and  $N$  are the associated multiplication operators on  $H$ . We then have

$$\widehat{MN^*}(x) = \widehat{M}(x)\overline{\widehat{N}(x)} = m(x)\overline{n(x)}.$$

Coburn showed that the metric  $\delta$  is a natural tool for studying the smoothness of Berezin transforms.

**Proposition 10** (Coburn [CO2]). If  $A$  is a bounded linear operator on  $H$ ,  $x, y \in X$  then

$$(7.3) \quad |\hat{A}(x) - \hat{A}(y)| \leq 2 \|A\| \delta(x, y).$$

Thus, also, if  $m \in M(H)$  and  $M$  is the associated multiplication operator then

$$(7.4) \quad |m(x) - m(y)| \leq 2 \|M\| \delta(x, y).$$

Estimate (7.3) is sharp in the sense that given  $H$ ,  $x$ , and  $y$  one can select  $A$  so that equality holds.

PROOF. It is standard that if  $A$  is bounded and  $T$  is trace class then  $AT$  is trace class and  $|\text{Trace}(AT)| \leq \|A\| \|T\|_{\text{Trace}}$ . Recall that  $P_x$  is the orthogonal projection onto the span of  $k_x$ . Direct computation shows  $\text{Trace}(AP_x) = \hat{A}(x)$ . Thus  $|\hat{A}(x) - \hat{A}(y)| \leq 2 \|A\| \|P_x - P_y\|_{\text{Trace}}$ . The proofs of (7.3) and (7.4) are then completed by taking note of Proposition 5. To see that the result is sharp evaluate both sides for the choice  $A = P_x - P_y$ . Details of those computation are in the proof of Proposition 5.  $\square$

**Remark 7.1.** By analysis of the two by two Pick matrix of  $M^*$  one sees that (7.4) is not sharp.

Suppose  $\gamma : (0, 1) \rightarrow X$  is a continuous curve in  $X$  and that  $f$  is a function defined in a neighborhood of the curve. We define the variation of  $f$  along the curve to be

$$\text{Var}(f) = \sup_{\gamma} \left\{ \sum_{i=1}^{n-1} |f(t_i) - f(t_{i+1})| : 0 < t_1 < \dots < t_n < 1, n = 1, 2, \dots \right\}.$$

**Corollary 1.** With  $H, X, \gamma, A$  as above:

$$\text{Var}(\hat{A}) \leq 2 \|A\| \ell_{\delta}(\gamma) = \|A\| \ell_{BS}(\gamma).$$

PROOF. If we start with a sum estimating the variation of  $\hat{A}$  and apply the previous proposition to each summand we obtain the first estimate. The second inequality follows from the first and Proposition 9.  $\square$

These issues are also studied when  $X$  has dimension greater than 1 and there is a rich relationship between the properties of Berezin transforms and the differential geometry associated with  $ds_{BS}^2$ , [CO], [CO2], [CL], [EZ], [EO], [BO].

**7.3. Limits Along Curves.** For the most commonly considered examples of a RKHS any curve in  $X$  which leaves every compact subset of  $X$  has infinite length when measured by any of the length functions we have been considering. However this is not always the case. For example, if  $X$  is the open unit disk and  $H$  is defined by the kernel function

$$K(z, y) = \frac{2 - z - \bar{y}}{1 - \bar{y}z}$$

then straightforward estimates show that along the positive axis

$$ds_{BS} = \left( \frac{1}{2\sqrt{1-r}} + o\left(\frac{1}{\sqrt{1-r}}\right) \right) dr;$$

hence the curve  $[0, 1)$  has finite length. For more discussion of this see [Mc]. This suggests there may be interesting limiting behavior as one traverses the curve. Suppose  $f$  is a function defined on  $\gamma$  except at the endpoints. Straightforward analysis then shows that if  $\text{Var}_\gamma(f) < \infty$  then  $f$  has limiting values along  $\gamma$  as one approaches the endpoints. Thus

**Corollary 2.** Given  $H, X$ ; suppose  $\gamma : [0, 1) \rightarrow X$  and  $\ell_{BS}(\gamma) < \infty$  then if  $m$  is any multiplier of  $H$  or, more generally, if  $\hat{A}$  is the Berezin transform of any bounded operator on  $H$ , then these functions have limits along  $\gamma$ ;

$$\exists \lim_{t \rightarrow 1} m(\gamma(t)); \exists \lim_{t \rightarrow 1} \hat{A}(\gamma(t)).$$

Furthermore there are choices of  $m$  or  $A$  for which the limits are not zero.

This invites speculation that something similar might be true for  $H$ , however that situation is more complicated. If we rescale the space then we do not change  $\delta$  and thus don't change the class of curves of finite length. However rescaling certainly can change the validity of statements that functions in the space have limits along certain curves. Thus the possibility of rescaling a space is an obstacle to having a result such as the previous corollary for functions in  $H$ . McCarthy showed in [Mc] that, in some circumstances, this is the only obstacle. We state a version of his result but will not include the proof.

**Proposition 11** (McCarthy [Mc]). Suppose  $H$  is a RKHS of holomorphic functions on the disk and that  $\gamma : [0, 1) \rightarrow \mathbb{D}$  and  $\ell_{BS}(\gamma) < \infty$ . There is a holomorphic function  $G$  such that every function in the rescaled space  $GH$  has a limit along  $\gamma$ . That is

$$\exists G, \forall h \in H, \exists \lim_{t \rightarrow 1} G(\gamma(t))h(\gamma(t)).$$

Furthermore  $G$  can be chosen so that for some  $h \in H$  this limit is not zero.

McCarthy's work was part of an investigation of an interface between operator theory and differential geometry that goes back (at least) to the work of Cowen and Douglas [CD]. They showed that one could associate to certain operators a domain in the plane and a Hermitian holomorphic line bundle in such a way that the domain and curvature of the line bundle formed a complete unitary invariant for the operator. That is, two such operators are unitarily equivalent if and only if the domains in the plane agree and the curvatures on the line bundles agree. An alternative presentation of their approach yields RKHSs of holomorphic functions which satisfy certain additional conditions. In this viewpoint the statement about unitary equivalence becomes the statement that for certain pairs of RKHSs of holomorphic functions, if their metrics  $\delta$  are the same then each space is a holomorphic rescaling of the other. The connection between the two viewpoints is that in terms of the kernel function of the RKHS, the curvature function at issue is

$$\mathcal{K}(z) = -\frac{\partial^2}{\partial z \partial \bar{z}} \log k_z(z).$$

Thus  $\mathcal{K}$  or, equivalently  $\delta$ , contains a large amount of operator theoretic information. However that information is not easy to access; which is why McCarthy's result is so nice and why these relations seem worth more study.

**7.4. RKHS's With Complete Nevanlinna Pick Kernels.** There is a class of RKHS's which are said to have complete Nevanlinna Pick kernels or complete NP kernels. The classical Hardy space is the simplest. The class is easy to define, we will do that in a moment, but the definition is not very informative. A great deal of work has been done in recent years studying this class of spaces. The book [AM] by Agler and McCarthy is a good source of information. In this and the next section we will see that for this special class of RKHS's the function  $\delta$  has additional properties.

We will say that the RKHS  $H$  has a complete NP kernel if there are functions  $\{b_i\}_{i=1}^\infty$  defined on  $X$  so that

$$1 - \frac{1}{K(x, y)} = \sum_{i=1}^{\infty} b_i(x) \overline{b_i(y)};$$

that is, if the function  $1 - 1/K$  is positive semidefinite.

Of the spaces in our earlier list of examples the Hardy spaces, generalized Dirichlet spaces, and the Dirichlet space have complete NP kernels. This is clear for the Hardy space, for the generalized Dirichlet spaces it follows from using the Taylor series for  $1/K$ , and for the Dirichlet space there is some subtlety involved in the verification. On the other hand neither the generalized Bergman spaces nor the Fock spaces have a complete NP kernel.

Suppose  $H$  is a RKHS of functions on  $X$  and  $x, y \in X$ ,  $x \neq y$ . Let  $G = G_{x,y}$  be the multiplier of  $H$  of norm 1 which has  $G(x) = 0$  and subject to those conditions maximizes  $\operatorname{Re} G(y)$ .

**Proposition 12.**

$$\operatorname{Re} G_{x,y}(y) \leq \delta(x, y).$$

If  $H$  has a complete NP kernel then

$$\operatorname{Re} G_{x,y}(y) = \delta(x, y).$$

and  $G$  is given uniquely by

$$G_{x,y}(\cdot) = \delta_H(x, y)^{-1} \left( 1 - \frac{k_y(x)k_x(\cdot)}{k_x(x)k_y(\cdot)} \right).$$

PROOF. [Sa, (5.9) pg 93]

□

**Remark 7.2.** The multipliers of  $H$  form a commutative Banach algebra  $\mathcal{M}$ . A classical metric on the spectrum of such an algebra, the Gleason metric, is given by

$$\delta_G(\alpha, \beta) = \{\sup \operatorname{Re} \alpha(M) : M \in \mathcal{M}, \|M\| = 1, \beta(M) = 0\},$$

[G], [L], [BW]. The points of  $X$  give rise to elements of the spectrum via  $\hat{x}(G) = G(x)$ . Thus if  $H$  has a complete NP kernel then  $\delta$  agrees with the Gleason metric:

$$\forall x, y \in X, \delta_G(\hat{x}, \hat{y}) = \delta(x, y).$$

**7.4.1. Generalized Blaschke Products.** If  $H = \mathcal{H}_1$ , the Hardy space, when we compute  $G$  we get

$$\begin{aligned}
G_{x,y}(z) &= \left| \frac{1 - \bar{x}y}{y - x} \right| \left( 1 - \frac{(1 - |x|^2)(1 - \bar{y}z)}{(1 - \bar{y}x)(1 - \bar{x}z)} \right) \\
&= \left| \frac{1 - \bar{x}y}{y - x} \right| \frac{\bar{y} - \bar{x}}{1 - \bar{y}x} \frac{z - x}{1 - \bar{x}z} \\
&= e^{i\theta} \frac{z - x}{1 - \bar{x}z}.
\end{aligned}$$

Thus  $G_{x,y}$  is a single Blaschke factor which vanishes at  $x$  and is normalized to be positive at the base point  $y$ .

Suppose now we have a RKHS  $H$  which has a complete NP kernel and let us suppose for convenience that it is a space of holomorphic functions on  $\mathbb{D}$ . Suppose we are given a set  $S = \{x_i\}_{i=1}^{\infty} \subset \mathbb{D}$  and we want to find a function in  $H$  and/or  $M(H)$  whose zero set is exactly  $S$ . We could use the functions  $G$  just described and imitate the construction of a general Blaschke product from the individual Blaschke factors. That is, pick  $x_0 \in \mathbb{D} \setminus S$  and consider the product

$$(7.5) \quad B(\zeta) = B_{S,x_0}(\zeta) = \prod_{i=1}^{\infty} G_{x_i,x_0}(\zeta).$$

If the product converges then  $B(\zeta)$  will be a multiplier of norm at most one and its zero set will be exactly  $S$ .

The multiplier norm dominates the supremum so the factors in (7.5) have modulus less than one. Hence the product either converges to a holomorphic function with zeros only at the points of  $S$  or the product diverges to the function which is identically zero. The same applies to the function  $B^2(\zeta)$ . We test the convergence of that product by evaluation at  $x_0$  and, recalling that  $G_{z,y}(y) = \delta_H(z, y)$ , we see

$$(7.6) \quad B_{S,x_0}^2(x_0) = \prod_{i=1}^{\infty} \delta^2(x_i, z_0),$$

$$(7.7) \quad = \prod_{i=1}^{\infty} \left( 1 - \frac{|k_{x_0}(x_i)|^2}{\|k_{x_0}\|^2 \|k_{x_i}\|^2} \right).$$

Recalling the conditions for absolute convergence of infinite products, we have established the following:

**Proposition 13.** The generalized Blaschke product  $B_{S,x_0}$  converges to an element of  $M(H)$  of norm at most one and with zero set exactly  $S$  if and only if the following two equivalent conditions hold

$$\begin{aligned}
&\prod_{i=1}^{\infty} \delta^2(x_i, x_0) > 0 \\
(7.8) \quad &\sum_{i=1}^{\infty} \frac{|k_{x_0}(x_i)|^2}{\|k_{x_0}\|^2 \|k_{x_i}\|^2} < \infty.
\end{aligned}$$

If the conditions do not hold then  $B_{S,x_0}$  is the zero function.

**Remark:** If  $1 \in H$  then  $M(H) \subset H$  and the proposition gives conditions that insure that there is a function in  $H$  with zero set exactly  $S$ .

**Corollary 3.** A sufficient condition for the set  $S$  to be a zero set for the Hardy space,  $H_1$ , or of  $M(H_1)$  which is known to be the space of bounded analytic functions in the disk, is that

$$(7.9) \quad \sum (1 - |x_i|^2) < \infty.$$

A sufficient condition for the set  $S$  to be a zero set for the Dirichlet space  $H_0$  or of  $M(H_0)$  is that

$$(7.10) \quad \sum \log \left( \frac{1}{1 - |x_i|^2} \right) < \infty.$$

PROOF. These are just the conclusions of the previous proposition applied to the Hardy space and the Dirichlet space with the choice of the origin for the basepoint.  $\square$

**Remark 7.3.** Condition (7.9) is the Blaschke condition which is well known to be necessary and sufficient for  $S$  to be the zero set of a function in  $\mathcal{H}_1$  or  $H^\infty$ .

The condition for  $S$  to be a zero set for the Dirichlet space was first given by Shapiro and Shields [SS] and the argument we gave descends from theirs. It is known that this condition is necessary and sufficient if  $S \subset (0, 1)$  but is not necessary in general.

**Remark 7.4.** In the next subsection we note that any  $H$  with a complete NP kernel is related to a special space of functions in a complex ball. Using that relationship one checks easily that the convergence criteria in the Proposition is a property of the set  $S$  and is independent of the choice of  $x_0$ .

*7.4.2. The Drury Arveson Hardy Space and Universal Realization.* For  $n = 1, 2, \dots$  we let  $\mathbb{B}^n$  denote the open unit ball in  $\mathbb{C}^n$ . We allow  $n = \infty$  and interpret  $\mathbb{B}^\infty$  to be the open unit ball of the one sided sequence space  $\ell^2(\mathbb{Z}_+)$ . For each  $n$  we define the  $n$ -dimensional Drury Arveson Hardy space,  $D_n$  to be the RKHS on  $\mathbb{B}^n$  with kernel function

$$K_n(x, y) = \frac{1}{1 - \langle x, y \rangle}.$$

Thus when  $n = 1$  we have the classical Hardy space.

For each  $n$  the kernel function  $K$  is a complete NP kernel. The spaces  $D_n$  are universal in the sense that any other RKHS with a complete NP kernel can be realized as a subspace of some  $D_n$ . If  $H$  is a RKHS on  $X$  and  $H$  has a complete NP kernel then there is for some  $n$ , possibly infinite, a mapping  $\gamma : X \rightarrow \mathbb{B}^n$  and a nonvanishing function  $b$  defined on  $X$  so that  $K_H$ , the kernel function for the space  $H$ , is given by

$$(7.11) \quad K_H(x, y) = b(x) \overline{b(y)} K_n(\gamma(x), \gamma(y)) = \frac{b(x) \overline{b(y)}}{1 - \langle \gamma(x), \gamma(y) \rangle}.$$

There is no claim of smoothness for  $\gamma$ . All this is presented in [AM].

The map  $\gamma$  can be used to pull back the pseudohyperbolic metric from  $\mathbb{B}^n$  to produce a metric on  $X$ . First we recall the basic facts about the pseudohyperbolic metric  $\mathbb{B}^n$ . Details about the construction of the metric and its properties can be found in [DW]; the discussion there is for finite  $n$  but the rudimentary pieces of theory we need for infinite  $n$  follow easily from the same considerations.

The pseudohyperbolic metric  $\rho$  on the unit disk,  $\mathbb{B}^1$ , can be described as follows. The disk possesses a transitive group of biholomorphic automorphisms,  $G = \{\phi_\alpha\}_{\alpha \in A}$ . Given a pair of points  $z, w \in \mathbb{B}^1$  select a  $\phi_\alpha \in G$  so that  $\phi_\alpha(z) = 0$ . The quantity  $|\phi_\alpha(w)|$  can be shown to be independent of the choice of  $\phi_\alpha$  and we define  $\rho(z, w) = |\phi_\alpha(w)|$ .

In this form the construction generalizes to  $\mathbb{B}^n$ . The  $n$ -ball has a transitive group of biholomorphic automorphisms,  $\mathcal{G} = \{\theta_\beta\}_{\beta \in B}$ . Given a pair of points  $z, w \in \mathbb{B}^n$  select a  $\theta_\beta \in \mathcal{G}$  so that  $\theta_\beta(z) = 0$ . The quantity  $|\theta_\beta(w)|$  can be shown to be independent of the choice of  $\theta_\beta$  and we define  $\rho_n(z, w) = |\theta_\beta(w)|$ . The only difference, and that is hidden by our notation, is that now  $|\cdot|$  denotes the Euclidean length of a vector rather than the modulus of a scalar. The function  $\rho_n$  can be shown to be a metric and to have the expected properties including invariance under  $\mathcal{G}$  and having an induced inner metric  $\rho_n^*$  that, up to a scalar factor, agrees with the distance induced by the Poincare-Bergman metric tensor. Particularly important for our purposes is that there is an analog of (2.3) [DW, pg 67]. For  $z, w \in \mathbb{B}^n$

$$(7.12) \quad 1 - \frac{(1 - |z|^2)(1 - |w|^2)}{|1 - \langle z, w \rangle|^2} = \rho_n^2(z, w).$$

An immediate consequence of the definition of  $\delta$ , the relationship (7.11), and the identity (7.12) is that the metric  $\delta$  on  $X$  is the pull back of  $\rho_n$  by  $\gamma$ . Put differently  $\gamma$  is an isometric map of  $(X, \delta_H)$  into  $(\mathbb{B}^n, \rho_{\mathbb{B}^n})$ . In particular the  $\delta$  metric on the Drury-Arveson space is the pseudohyperbolic metric on  $\mathbb{B}^n$ :  $\delta_{D_n} = \rho_{\mathbb{B}^n}$ .

## 8. Invariant Subspaces and Their Complements

Suppose we are given RKHSs on a set  $X$  and linear maps between them. We would like to use the  $\delta$ s on  $X$  to study the relation between the function spaces and to study the linear maps. The goal is broad and vague. Here we just report on a few very special cases.

We will consider a RKHS  $H$  of functions on a set  $X$ , a closed multiplier invariant subspace  $J$  of  $H$ ; that is we require that if  $j \in J$  and  $m$  is a multiplier of  $H$  then  $mj \in J$ . We will also consider the  $J^\perp$ , the orthogonal complement of  $J$ . The spaces  $J$  and  $J^\perp$  are RKHSs on  $X$  and we will be interested in the relationship between the metrics  $\delta_H$ ,  $\delta_J$ , and  $\delta_{J^\perp}$ . Because we are working with a subspace and its orthogonal complement there is a simple relation between the kernel functions. Let  $\{k_x\}$  be the kernel function  $H$ ,  $\{j_x\}$  those for  $J$  and  $\{j_x^\perp\}$  those for  $J^\perp$ . We then have,  $\forall x \in X$

$$(8.1) \quad k_x = j_x + j_x^\perp.$$

In terms of  $P$ , the orthogonal projection of  $H$  onto  $J$ , and  $P^\perp = I - P$ , we have  $j_x = Pk_x$ ,  $j_x^\perp = P^\perp k_x$ .

**8.1. The Hardy Space.** We begin with  $\mathcal{H}_1$ , the Hardy space. In that case there is a good description of the invariant subspaces, the computations go smoothly, and the resulting formulas are simple. If  $J$  is an invariant subspace then there is an inner function  $\Theta_J$  so that  $J = \Theta_J \mathcal{H}_1$ . The kernel functions are, for  $z, w \in \mathbb{D}$ ,  $\Theta_J(z) \neq 0$  are given by

$$j_z(w) = \frac{\overline{\Theta_J(z)\Theta_J(w)}}{1 - \bar{z}w}.$$

Thus if  $\Theta_J(z)\Theta_J(z') \neq 0$  then

$$\delta_J(z, z') = \delta_{\mathcal{H}_1}(z, z').$$

In the other cases, by our convention,  $\delta_J$  is undefined.

Taking into account the formula for  $j_z$  and (8.1) we find

$$j_z^\perp(w) = \frac{1 - \overline{\Theta_J(z)}\Theta_J(w)}{1 - \bar{z}w}$$

and hence

$$1 - \delta_{J^\perp}^2(z, w) = \frac{(1 - |\Theta_J(z)|^2)(1 - |\Theta_J(w)|^2)}{\left|1 - \overline{\Theta_J(z)}\Theta_J(w)\right|^2} \frac{(1 - |z|^2)(1 - |w|^2)}{|1 - \bar{z}w|^2}.$$

We can now use (2.3) on both fractions and continue with

$$1 - \delta_{J^\perp}^2(z, w) = \frac{1 - \rho^2(z, w)}{1 - \rho^2(\Theta_J(z), \Theta_J(w))}.$$

Doing the algebra we obtain

$$\delta_{J^\perp}(z, w) = \sqrt{\frac{\rho^2(z, w) - \rho^2(\Theta_J(z), \Theta_J(w))}{1 - \rho^2(\Theta_J(z), \Theta_J(w))}}.$$

In particular

$$\delta_{J^\perp} \leq \rho = \delta_{\mathcal{H}_1}$$

with equality holding if and only if  $\Theta_J(z) = \Theta_J(w)$ .

**8.2. Triples of Points and the Shape Invariant.** When we move away from the Hardy space computation becomes complicated. Suppose  $H$  is a RKHS on  $X$  with kernel functions  $\{k_x\}$  and associated distance function  $\delta$ . Select distinct  $x, y, z \in X$ . We consider the invariant subspace  $J$  of functions which vanish at  $x$ , the orthogonal complement of the span of  $k_x$ . We will denote the kernel functions for  $J$  by  $\{j_z\}$ . We want to compute  $\delta_J(y, z)$  in terms of other data.

For any  $\zeta \in X$ ,  $j_z$  equals  $k_\zeta$  minus the projection onto  $J^\perp$  of  $k_\zeta$ . The space  $J^\perp$  is one dimensional and spanned by  $k_x$  hence we can compute explicitly

$$j_y = k_y - \frac{k_y(x)}{\|k_x\|^2} k_x,$$

and there is a similar formula for  $j_z$ . We will need  $|\langle j_y, j_z \rangle|^2$ .

$$\begin{aligned} |\langle j_y, j_z \rangle|^2 &= \left| k_y(z) - \frac{k_y(x)}{\|k_x\|^2} k_x(z) \right|^2 \\ (8.2) \quad &= |k_y(z)|^2 + \frac{|k_y(x)k_x(z)|^2}{\|k_x\|^4} - 2 \operatorname{Re} k_y(z) \frac{\overline{k_y(x)}}{\|k_x\|^2} k_x(z). \end{aligned}$$

Recall that for any two distinct elements  $\alpha, \beta$  of the set  $\{x, y, z\}$  we have

$$(8.3) \quad |k_\alpha(\beta)|^2 = |k_\beta(\alpha)|^2 = \|k_\alpha\|^2 \|k_\beta\|^2 (1 - \delta^2(\alpha, \beta)).$$

We use this in (8.2) to replace the quantities such as the ones on the left in (8.3) with the one on the right. Also, we will write  $\delta_{\alpha\beta}^2$  rather than  $\delta^2(\alpha, \beta)$ . For the same  $\alpha, \beta$  we define  $\theta_{\alpha\beta}$  and  $\phi_{\alpha\beta}$  with  $0 \leq \theta_{\alpha\beta} \leq \pi$ , and  $\phi_{\alpha\beta}$  with  $0 \leq \phi_{\alpha\beta} < \pi$  by

$$k_\alpha(\beta) = \langle k_\alpha, k_\beta \rangle = \|k_\alpha\| \|k_\beta\| (\cos \theta_{\alpha\beta}) e^{i\phi_{\alpha\beta}}$$

and we set

$$\Upsilon = \cos \theta_{xy} \cos \theta_{yz} \cos \theta_{zx} \cos(\phi_{xy} + \phi_{yz} + \phi_{zx}).$$

We now continue from (8.2) with

$$\begin{aligned} |\langle j_y, j_z \rangle|^2 &= \|k_z\|^2 \|k_y\|^2 (1 - \delta_{zy}^2) + \|k_y\|^2 \|k_z\|^2 (1 - \delta_{xy}^2) (1 - \delta_{xz}^2) - 2 \|k_y\|^2 \|k_z\|^2 \Upsilon \\ &= \|k_z\|^2 \|k_y\|^2 \{(1 - \delta_{zy}^2) + (1 - \delta_{xy}^2) (1 - \delta_{xz}^2) - 2\Upsilon\}. \end{aligned}$$

Similar calculations give

$$\begin{aligned} \|j_y\|^2 &= \|k_y\|^2 - \frac{|k_x(y)|^2}{\|k_x\|^2} = \|k_y\|^2 - \|k_y\|^2 (1 - \delta_{xy}^2) = \|k_y\|^2 \delta_{xy}^2 \\ \|j_z\|^2 &= \|k_z\|^2 \delta_{xz}^2 \end{aligned}$$

Hence

$$\begin{aligned} \delta_J^2(y, z) &= 1 - \frac{|\langle j_y, j_z \rangle|^2}{\|j_y\|^2 \|j_z\|^2} \\ &= 1 - \frac{\|k_z\|^2 \|k_y\|^2 \{(1 - \delta_{zy}^2) + (1 - \delta_{xy}^2) (1 - \delta_{xz}^2) - 2\Upsilon\}}{\|k_y\|^2 \delta_{xy}^2 \|k_z\|^2 \delta_{xz}^2} \\ &= \frac{\delta_{xy}^2 \delta_{xz}^2 - (1 - \delta_{zy}^2) - (1 - \delta_{xy}^2) (1 - \delta_{xz}^2) + 2\Upsilon}{\delta_{xy}^2 \delta_{xz}^2} \\ &= \frac{\delta_{xy}^2 + \delta_{xz}^2 + \delta_{zy}^2 - 2 + 2\Upsilon}{\delta_{xy}^2 \delta_{xz}^2}. \end{aligned}$$

Thus we can write the very symmetric formula

$$(8.4) \quad \delta(y, z) \delta_J(y, z) = \frac{\sqrt{\delta_{xy}^2 + \delta_{xz}^2 + \delta_{zy}^2 - 2 + 2\Upsilon}}{\delta_{xy} \delta_{xz} \delta_{yz}}.$$

One reason for carrying this computation through is to note the appearance of  $\Upsilon$ . This quantity, which is determined by the ordered triple  $\{k_x, k_y, k_z\}$  and which is invariant under cyclic permutation of the three, is a classical invariant of projective and hyperbolic geometry. It is called the shape invariant. The triple determines an ordered set of three points in  $P(H)$  the projective space over  $H$ ;  $p_x = [k_x]$ ,  $p_y$ , and  $p_z$ . Modulo some minor technicalities which we omit, one can regard the three as vertices of an oriented triangle  $T_{xyz}$ . The edges of the triangle are the geodesics connecting the vertices and the surface is formed by the collection of all geodesics connecting points on the edges. In Euclidian space two triangles are congruent, one can be moved to the other by an action of the natural isometry group, if and only if the set of side lengths agree. That is not true in projective space. The correct statement there is that two triangles are congruent if and only if the three side lengths and the shape invariants match [BT], [BR]. Using the natural geometric structure of complex projective space one can also define and compute the area of  $T_{xyz}$ . It turns out, roughly, that once the side lengths are fixed then  $\Upsilon$  determines

the area of  $T_{xyz}$  and vice versa [HM]. Further discussion of the shape invariant is in [Go] and [BS].

The reason for mentioning all this is that  $\Upsilon$  was the one new term that appeared in (8.4) and it is slightly complicated. The fact that this quantity has a life of its own in geometry suggests that perhaps the computations we are doing are somewhat natural and may lead somewhere interesting.

In Bøe's work on interpolating sequences in RKHS with complete NP kernel [B] (see also [Sa]) he makes computations similar in spirit and detail to the ones above. Informally, he is extracting analytic information from a geometric hypothesis. It is plausible that knowing how such techniques could be extended from three points to  $n$  points would allow substantial extension of Bøe's results.

**8.3. Monotonicity Properties.** We saw that when  $J$  was an invariant subspace of the Hardy space  $\mathcal{H}_1$  then for all  $x, y$  in the disk

$$\delta_J(x, y) = \delta_{\mathcal{H}_1}(x, y) \geq \delta_{J^\perp}(x, y).$$

It is not clear what, if any, general pattern or patterns this is an instance of. Here we give some observations and computations related to that question.

8.3.1. *Maximal Multipliers.* Fix  $H$  and  $X$ . Recall Proposition 12; given  $x, y \in X$ ,  $x \neq y$  we denote by  $G_{x,y}$  be the multiplier of  $H$  of norm 1 which has  $G(x) = 0$  and subject to those conditions maximizes  $\operatorname{Re} G(y)$ . The Proposition stated that  $\operatorname{Re} G_{x,y}(y) \leq \delta_H(x, y)$  and that equality sometimes held. If equality does hold we will say that  $x, y$  have an maximal multiplier and we will call  $G_{x,y}$  the maximal multiplier.

**Proposition 14.** If  $x, y \in X$  have a maximal multiplier and  $J$  is any closed multiplier invariant subspace of  $H$  then  $\delta_J(x, y) \geq \delta_H(x, y)$ .

PROOF. Because  $J$  is a closed multiplier invariant subspace of  $H$ , the maximal multiplier  $G_{x,y}$  is also a multiplier of  $J$  and has a norm, as a multiplier on  $J$ , at most one. Thus  $G_{x,y}$  is a competitor in the extremal problem associated with applying Proposition 12 to  $J$ . Hence, by that proposition we have  $\operatorname{Re} G_{x,y}(y) \leq \delta_J(x, y)$ . On the other hand our hypothesis is that  $\operatorname{Re} G_{x,y}(y) = \delta_H(x, y)$ .  $\square$

If  $H$  has a complete NP kernel then every pair of points,  $x, y$ , has a maximal multiplier. Hence

**Corollary 4.** If  $H$  has a complete NP kernel and  $J$  is any closed multiplier invariant subspace of  $H$  then for all  $x, y \in X$ ,  $\delta_J(x, y) \geq \delta_H(x, y)$ .

The converse of the corollary is not true. Having a complete NP kernel is not a necessary condition in order for every pair of points to have an maximal multiplier; it is sufficient that the kernel have the scalar two point Pick property, see [AM, Ch. 6,7].

8.3.2. *Spaces with Complete Nevanlinna Pick Kernels.* Suppose that  $H$  is a RKHS on  $X$  with a complete NP kernel  $K(\cdot, \cdot)$ . Suppose also, and this is for convenience, that we have a distinguished point  $\omega \in X$  such that  $\forall x \in X$ ,  $K(\omega, x) = K(x, \omega) = 1$ . The following information about invariant subspaces of  $H$  is due to McCullough and Trent [MT], further information is in [GRS].

**Proposition 15.** Suppose  $J$  is a closed multiplier invariant subspace of  $H$ . There are multipliers  $\{m_i\}$  so that the reproducing kernel for  $J$  is

$$(8.5) \quad K_J(x, y) = \left( \sum m_i(x) \overline{m_i(y)} \right) K(x, y).$$

**Corollary 5.** If  $H$  has a complete NP kernel and  $J$  is any closed multiplier invariant subspace of  $H$  then for all  $x, y \in X$

$$\delta_J(x, y) \geq \delta_H(x, y) \geq \delta_{J^\perp}(x, y).$$

PROOF. We start with formula (8.5) which we rewrite for convenience as

$$(8.6) \quad K_J(x, y) = A(x, y) K(x, y).$$

The first inequality is the statement of the previous corollary. Alternatively we could start from the previous equality and use (5.1) to compare  $\delta_H$  and  $\delta_J$  yielding a quantitative version of the desired inequality.

For the second inequality first note that  $K_{J^\perp} = K - K_J = K - AK = (1 - A)K$ .

(Note that for any  $x$ ,  $A(x, x) \leq 1$  because it is the ratio of the squared norms of two kernel functions and the one on top, being a projection of the one on bottom, has smaller norm. Also, by Cauchy-Schwarz,  $|A(x, y)|^2 \leq A(x, x)A(y, y) \leq 1$ . To rule out the case of equality note that if  $A(x, x) = 1$  then  $k_x \in J$  and hence every function in  $J^\perp$  vanishes at  $x$  which puts  $x$  outside the domain of  $\delta_{J^\perp}$ .)

Recalling the formula for  $\delta$  we see that our claim will be established if we can show for the  $x, y \in X$  that are covered by the claim we have

$$\frac{|1 - A(x, y)|^2}{(1 - A(x, x))(1 - A(y, y))} \geq 1.$$

We have

$$\begin{aligned} |1 - A(x, y)|^2 &\geq (1 - |A(x, y)|)^2 \\ &\geq \left(1 - A(x, x)^{1/2} A(y, y)^{1/2}\right)^2 \\ &= 1 - 2A(x, x)^{1/2} A(y, y)^{1/2} + A(x, x)A(y, y) \\ &\geq 1 - A(x, x) - A(y, y) + A(x, x)A(y, y) \\ &= (1 - A(x, x))(1 - A(y, y)). \end{aligned}$$

Here the passage from the first line to the second uses Cauchy-Schwarz, the passage from third to fourth uses the arithmetic mean, geometric mean inequality.  $\square$

**8.3.3. Inequalities in the Other Direction; Bergman Type Spaces.** Let  $H = \mathcal{H}_2$  be the Bergman space. That is,  $H$  is the RKHS of holomorphic functions on the disk with reproducing kernel  $K(z, w) = (1 - \bar{w}z)^{-2}$ . Let  $J$  be the invariant subspace consisting of all functions in  $H$  which vanish at the origin.

**Proposition 16.** For all  $z, w \in \mathbb{D}$

$$\delta_J(z, w) \leq \delta_H(z, w).$$

PROOF. We have

$$\begin{aligned} K_J(z, w) &= K(z, w) - 1 = \frac{1}{(1 - \bar{w}z)^2} - 1 \\ &= \frac{\bar{w}z(2 - \bar{w}z)}{(1 - \bar{w}z)^2} = \frac{2\bar{w}z(1 - \frac{\bar{w}z}{2})}{(1 - \bar{w}z)^2} \\ &= B(z, w)K(z, w). \end{aligned}$$

At this stage we can see the difference between this situation and the one in the previous section. Here the ratio  $K_J/K$  is *not* a positive definite function. To finish we need to show

$$\frac{|B(z, w)|^2}{B(z, z)B(w, w)} \geq 1.$$

Thus we need to show

$$|2 - \bar{w}z|^2 \geq (2 - \bar{w}w)(2 - \bar{z}z).$$

Equivalently, we need

$$-4 \operatorname{Re} \bar{w}z \geq -2\bar{w}w - 2\bar{z}z.$$

This follows from the inequality between the arithmetic and geometric means.  $\square$

In fact this example is just the simplest case of a general pattern introduced in [HJS] and [MR]. In [MR] McCullough and Richter introduce a general class of RKHS which share many of the properties of the Bergman space. In particular their work covers the spaces  $\mathcal{H}_\alpha$ ,  $1 \leq \alpha \leq 2$ , and we will focus on that case. Suppose  $V$  is an invariant subspace of some  $\mathcal{H}_\alpha$ ,  $1 \leq \alpha \leq 2$  and that  $V$  has index 1, that is  $\dim V \ominus zV = 1$ . Let  $\{k_z\}$  be the reproducing kernels for  $\mathcal{H}_\alpha$  and  $\{j_z\}$  be those for  $V$ .

**Proposition 17** (Corollary 0.8 of [MR]). There is a function  $G \in \mathcal{H}_\alpha$  and a positive semidefinite sesquianalytic function  $A(z, w)$  so that for  $z, w \in \mathbb{D}$

$$j_z(w) = \overline{G(z)}G(w)(1 - \bar{z}wA(z, w))k_z(w).$$

**Corollary 6.**

$$\delta_V(z, w) \leq \delta_{\mathcal{H}_\alpha}(z, w).$$

PROOF. The factors of  $G$  do not affect  $\delta$ . After they are dropped the argument is then the same as in the proof of Corollary 5.

$$\begin{aligned} |1 - \bar{z}wA(z, w)|^2 &\geq (1 - |\bar{z}wA(z, w)|)^2 \\ &\geq (1 - |z|A(z, z)^{1/2}|w|A(w, w)^{1/2})^2 \\ &\geq 1 - 2|z|A(z, z)^{1/2}|w|A(w, w)^{1/2} + |z|^2A(z, z)^{1/2}|w|^2A(w, w) \\ &\geq 1 - |z|^2A(z, z) - |w|^2A(w, w) + |z|^2A(z, z)|w|^2A(w, w) \\ &\geq (1 - |z|^2A(z, z))(1 - |w|^2A(w, w)) \end{aligned}$$

which is what is needed.

Again, the passage from the first line to the second uses Cauchy-Schwarz, the passage from third to fourth uses the arithmetic mean, geometric mean inequality.  $\square$

It is not clear in this context what happens with spaces of the form  $J^\perp$ . The following computational example suggests the story may be complicated. Let  $J$  be the invariant subspace of  $H = \mathcal{H}_2$  consisting of functions  $f$  with  $f(0) = f'(0) = 0$ . The reproducing kernel for  $J^\perp$  is

$$K_{J^\perp}(z, w) = 1 + 2\bar{w}z.$$

To compare  $\delta_{J^\perp}$  with  $\delta_H$  we compare the quantities  $1 - \delta^2$ . If we are looking for an example where the results for spaces with complete NP kernels fails then we need to find  $z, w$  such that

$$\frac{|1 + 2\bar{w}z|^2}{(1 + 2|w|^2)(1 + 2|z|^2)} \leq \frac{(1 - |w|^2)^2 (1 - |z|^2)^2}{|1 - \bar{w}z|^4}.$$

This certainly does not hold if either  $|z|$  or  $|w|$  is close to 1 however things are different near the origin. Suppose  $z = -w = t > 0$ . The inequality we want is

$$\frac{(1 - 2t^2)^2}{(1 + 2t^2)^2} \leq \frac{(1 - t^2)^4}{(1 + t^2)^4}.$$

This fails if  $t$  is close to one but we study it now for  $t$  near 0. In that case the left hand side is  $1 - 8t^2 + 32t^4 - 96t^6 + O(t^8)$  and the other side is  $1 - 8t^2 + 32t^4 - 88t^6 + O(t^8)$ . Hence for small  $t$  the inequality does hold.

**Proposition 18.** With  $H$  and  $J$  as described, if  $t$  is small and positive then

$$\delta_{J^\perp}(t, -t) > \delta_H(t, -t).$$

For  $t$  near 1 the inequality is reversed.

## 9. Questions

We mentioned in the introduction that most questions in this area have not been studied. Here we mention a few specific questions which had our interest while preparing this paper and which indicate how little is known.

1. Suppose  $H$  and  $H'$  are two RKHSs on the same  $X$  with distance functions  $\delta$  and  $\delta'$ ; and suppose further, in fact, that  $H$  and  $H'$  are the same spaces of functions with equivalent norms. What conclusions follow about  $\delta$  and  $\delta'$ ?
2. In the other direction, what conclusion can one draw about the relation between  $H$  and  $H'$  if the identity map from  $(X, \delta)$  to  $(X, \delta')$  is, say, a contraction or is bilipschitz?
3. It seems plausible that there is a more complete story to be told related to Corollary 5. What is the full class of RKHS for which those conclusions hold? What assumptions beyond those conclusions are needed to insure that the space being considered has a complete NP kernel?
4. Given  $X$  what metrics  $\delta$  can arise from a RKHS  $H$  on  $X$ . The question is extremely broad but notice that if you assume further that  $H$  must have a complete NP kernel then, by virtue of the realization theorem, a necessary and sufficient condition is that for some  $n$  there is an isometric map of  $(X, \delta)$  into  $(\mathbb{B}^n, \rho)$ . Although that answer is perhaps not particularly intuitive it does give a condition that is purely geometric, none of the Hilbert space discussion survives.

### References

- [AM] Agler, Jim; McCarthy, John E. *Pick interpolation and Hilbert function spaces*. Graduate Studies in Mathematics, 44. American Mathematical Society, Providence, RI, 2002.
- [AE] Ali, S. Twareque; Engliš, Miroslav *Quantization methods: a guide for physicists and analysts*. Rev. Math. Phys. 17 (2005), no. 4, 391–490.
- [BW] Bear, H. S.; Weiss, Max L. *An intrinsic metric for parts*. Proc. Amer. Math. Soc. 18 1967 812–817.
- [BS] Berceanu, Stefan; Schlichenmaier, Martin *Coherent state embeddings, polar divisors and Cauchy formulas*. J. Geom. Phys. 34 (2000), no. 3-4, 336–358.
- [BT] Blaschke, Wilhelm; Terheggen, Hans, *Trigonometria hermitiana*. Rend. Sem. Mat. Roma 3, (1939). 153–161.
- [B] Bøe, Bjarte *An interpolation theorem for Hilbert spaces with Nevanlinna-Pick kernel*. Proc. Amer. Math. Soc. 133 (2005), no. 7, 2077–2081.
- [BO] Bommier-Hato, Lipschitz estimates for the Berezin transform, *J. Funct. Spaces Appl.* to appear.
- [BR] Brehm, Ulrich *The shape invariant of triangles and trigonometry in two-point homogeneous spaces*. Geom. Dedicata 33 (1990), no. 1, 59–76.
- [CO] Coburn, L. A. *A Lipschitz estimate for Berezin's operator calculus*. Proc. Amer. Math. Soc. 133 (2005), no. 1, 127–131.
- [CO2] Coburn, L. A. *Sharp Berezin Lipschitz estimates*. Proc. Amer. Math. Soc. 135 (2007), no. 4, 1163–1168.
- [CL] Coburn, L. A.; Li, Bo *Directional derivative estimates for Berezin's operator calculus*. Proc. Amer. Math. Soc. 136 (2008), no. 2, 641–649.
- [CD] Cowen, M. J.; Douglas, R. G. *Complex geometry and operator theory*. Acta Math. 141 (1978), no. 3-4, 187–261.
- [DW] Duren, Peter; Weir, Rachel *The pseudohyperbolic metric and Bergman spaces in the ball*. Trans. Amer. Math. Soc. 359 (2007), no. 1, 63–76.
- [E] Engliš, Miroslav *Weighted Bergman kernels and quantization*. Comm. Math. Phys. 227 (2002), no. 2, 211–241.
- [EO] Engliš, Miroslav; Otáhalová, Renata *Covariant Derivatives of the Berezin Transform*, preprint, 2010.
- [EZ] Engliš, Miroslav; Zhang, Genkai *On the derivatives of the Berezin transform*. Proc. Amer. Math. Soc. 134 (2006), no. 8, 2285–2294.
- [G] Gleason, Andrew *Seminars on analytic functions*, Vol. 2, pp. 213–226; Princeton Univ. Press, Princeton, N.J., 1957.
- [Go] Goldman, William M. *Complex hyperbolic geometry*. Oxford Mathematical Monographs. Oxford Science Publications. 1999.
- [GRS] Greene, Devin C. V. Richter, Stefan; Sundberg, Carl *The structure of inner multipliers on spaces with complete Nevanlinna-Pick kernels*. J. Funct. Anal. 194 (2002), no. 2, 311–331.
- [HM] Hangan, Th.; Masala, G. *A geometrical interpretation of the shape invariant for geodesic triangles in complex projective spaces*. Geom. Dedicata 49 (1994), no. 2, 129–134.
- [HJS] Hedenmalm, Håkan; Jakobsson, Stefan; Shimorin, Sergei *A biharmonic maximum principle for hyperbolic surfaces*. J. Reine Angew. Math. 550 (2002), 25–75.
- [JPR] Janson, Svante; Peetre, Jaak; Rochberg, Richard *Hankel forms and the Fock space*. Rev. Mat. Iberoamericana 3 (1987), no. 1, 61–138.
- [JP] Jarnicki, Marek; Pflug, Peter *Invariant distances and metrics in complex analysis*. de Gruyter Expositions in Mathematics, 9. Walter de Gruyter & Co., Berlin, 1993.
- [K] Kobayashi, Shoshichi *Geometry of bounded domains*. Trans. Amer. Math. Soc. 92 1959 267–290.
- [L] Lewittes, Joseph *A note on parts and hyperbolic geometry*. Proc. Amer. Math. Soc. 17 1966 1087–1090.
- [MPS] Mazur, T.; Pflug, P.; Skwarczyński, M. *Invariant distances related to the Bergman function*. Proc. Amer. Math. Soc. 94 (1985), no. 1, 72–76.
- [Mc] McCarthy, John E. *Boundary values and Cowen-Douglas curvature*. J. Funct. Anal. 137 (1996), no. 1, 1–18.
- [MR] McCullough, Scott; Richter, Stefan *Bergman-type reproducing kernels, contractive divisors, and dilations*. J. Funct. Anal. 190 (2002), no. 2, 447–480.

- [MT] McCullough, Scott; Trent, Tavan T. Invariant subspaces and Nevanlinna-Pick kernels. *J. Funct. Anal.* 178 (2000), no. 1, 226–249.
- [O1] Odzijewicz, Anatol On reproducing kernels and quantization of states. *Comm. Math. Phys.* 114 (1988), no. 4, 577–597.
- [O2] Odzijewicz, Anatol Coherent states and geometric quantization. *Comm. Math. Phys.* 150 (1992), no. 2, 385–413.
- [PV] Provost, J. P.; Vallee, G. Riemannian structure on manifolds of quantum states. *Comm. Math. Phys.* 76 (1980), no. 3, 289–301.
- [Sa] Sawyer, Eric T. Function theory: interpolation and corona problems. Fields Institute Monographs, 25. AMS, Providence, RI; Fields Institute Toronto, ON, 2009.
- [SS] Shapiro, H. S.; Shields, A. L. On the zeros of functions with finite Dirichlet integral and some related function spaces. *Math. Z.* 80 1962 217–229.
- [S] Skwarczyński, M. The invariant distance in the theory of pseudoconformal transformations and the Lu Qi-keng conjecture. *Proc. Amer. Math. Soc.* 22 1969 305–310.

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI BOLOGNA, 40127 BOLOGNA, ITALY  
*E-mail address:* `arcozzi@dm.unibo.it`

DEPARTMENT OF MATHEMATICS, WASHINGTON UNIVERSITY, ST. LOUIS, MO 63130, U.S.A  
*E-mail address:* `rr@math.wustl.edu`

DEPARTMENT OF MATHEMATICS & STATISTICS, McMaster UNIVERSITY, HAMILTON, ONTARIO, L8S 4K1, CANADA  
*E-mail address:* `sawyer@mcmaster.ca`

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, 686 CHERRY STREET, ATLANTA, GA USA 30332-0160  
*E-mail address:* `wick@math.gatech.edu`